



TRUE-AI 교육연구단 세미나

CoMet: Context and Multiplicity Decomposition for Multimodal Uncertainty Estimation

전상혁

Princeton University

2026.10.20.(화요일) 16:00-17:00 ECC B161호

Abstract

Uncertainty estimation has been a long-standing challenge in AI models; it amounts to "knowing what you don't know," and metacognition is notoriously difficult even for humans (cf. the Dunning-Kruger effect). Although it is still far from solved even in simpler classification systems, tackling it in multimodal large language models (MLLMs) is becoming increasingly important. Within MLLMs, uncertainty can stem from any of the diverse sources as well as from their relationships, and further can stem from the unbounded answers in the open-ended setting. To tackle the issues, we propose CoMet, an MLLM uncertainty estimation method by decomposing uncertainty into a context-specific term and a multiplicity-specific term. The former captures ambiguity induced by the given context (e.g., task or prompt), while the latter captures how many plausible answers determined by the context remain compatible with the given input. We train a lightweight post-hoc uncertainty module to estimate these quantities, which enables efficient uncertainty estimation without autoregressive answer generation or repeated sampling. Experiments on various open-ended multimodal benchmarks, hallucination detection, and multiple-choice visual question answering benchmarks show that CoMet consistently improves uncertainty estimation over existing baselines while remaining efficient in practice.

Biography

2025.07 – present Ph.D., Princeton University
-2025.06 NAVER AI Lab
-2018.01 Kakao Corp
-2016.01 M.S. in KAIST
-2014.02 B.S. in KAIST

· 주최/주관: 4단계 BK21 사업 고신뢰 고효율 인공지능 교육연구단
(Trustworthy and Resource-efficient Unified Evolving AI)
G-LAMP 인간중심인공지능연구원 (이지영 교수/인공지능)

